



International Journal of Computer in Humanities



Refining Abstrak using Machine Translation

Tatan Tawami¹, Susi Yuliawati²

¹Departemen Sastra Inggris, Universitas Komputer Indonesia, Bandung, Indonesia

²Departemen Sastra Inggris, Universitas Padjadjaran, Bandung, Indonesia

Correspondence: E-mail: tatan.tawami@email.unikom.ac.id

ABSTRACTS

This paper investigates steps in which abstrak of a paper may be well translated into abstract using particular computer aided tool, the machine translation. In analysing the steps mentioned, the narrative-descriptive method is implemented in order to specify the sequence of the steps. In the process, it turned out that there were at least 4 steps to perform in order to create a well translated abstrak into abstract. Four steps can be done in order to create one well-translated abstrak. Each steps requires systematic analysis in order to achieve correct and accurate translation. Finishing the steps mean the translated abstrak may require further action in order to achieve best result. By performing these systematic and step by step sequences, the abstrak can express your best thought, additionally, the steps can even be extended to the entire paper to achieve the accurate translation.

ARTICLE INFO

Article History:

Keywords:

*Translation,
Computer-Aided Tool*

I. INTRODUCTION

Translating abstrak (Bahasa) of a paper into abstract (English) can easily be done with web-based machine translation (MT) (Groves *et al.*, 2015) such as Google Translate. However, translating it into the well-formed one using the same computer-aided tool can be tricky in certain ways because the tool does not always provide best and natural translation as expected. This happens because google translate cannot automatically detect the dynamic use of the source language of the user.

It is noticeably acknowledged that Google Translate is a system that calculates source language (SL) expression or phrases--not words--into its best target language (TL) form by continuously absorbing the information into its database (Groves *et al.*, 2015). However, as a machine, Google Translate cannot easily transform the sophisticated and indigenous SL expression and phrases (Groves *et al.*, 2015). Additionally, despite its rapidity in translating SL into TL, Google Translate is said to be less accurate (ElShiekh *et al.*, 2012) than the human translation in terms of understanding the cases in each linguistic level.

Research on this issue has been conducted, specifically in the translation process and possible refining process for the translation by investigating the errors in the linguistic levels and the cultural issues (ElShiekh *et al.*, 2012). Other related research discusses the coding languages used in the MT, the use of various models and rules in the implementation, and the sentiment analysis as well (Balahur *et al.*, 2012). In addition to that, there is also research investigating the translation efficiency of Google Translate compared

to human translation (Gracia *et al.*, 2009) through quantitative analysis (Afshin *et al.*, 2009) and qualitative analysis (Läubli *et al.*, 2017). Last but not least, research in the errors made in Google Translate has also been conducted (Ismail *et al.*, 2016); it discusses the categories in which the errors occurred.

However, the previous researches do not specifically discuss the issue of how lexical choice is used well in both cultural and natural language equivalents. It, again, does not describe the steps for refining the lexical choice so that the naturalness of the translation can be achieved. Additionally, what options are there available for refining the translation are not clearly mentioned (Raharjo, 2017). This research attempts to offer the available translation refining tools in its simplest form.

As clearly suggested, this research seeks to discover what steps can be taken to refine the translation using the existing machine translation and computer, at the same time it offers the simplest form for translation refining. To support this goal, the descriptive-analytic method is implemented so that each step shall attempt to analyse and describe the way through which refinement is done. This is believed to provide readers with a step-by-step process for improving their translation using Google Translate.

2. METHODOLOGY

This research is a prospective case study (Raharjo, 2017). using the analytic-descriptive method (Oshima *et al.*, 1997). Prospective case study means that the data analysed is intended to provide better ways in translating the source language into the target language, which later will result in better outcomes. The

data, specifically the abstrak, was a self-documented work for an individual client on March 16, 2019. The selection of this abstract was due to its varied cases for translation, especially when using google translate. Each sentence of the abstrak was then analysed for particular idiomatic expressions or sophisticated phrases and other elements constructing the sentence and paragraph coherence. Every step of the analysis was described systematically in order to obtain a better description concerning the case under the study.

3. RESULTS AND DISCUSSION

In order to refine the translation result, several steps can be taken into account: (a) correcting the source language so that Google Translate can identify the intended meaning, (b) do a sentence by sentence translation, or phrase by phrase when necessary, (c) when finding odd expression, consult dictionary for its use, (d) do final checking to the translation. Therefore, it is beneficial to show the way we treat the source language in order to achieve better machine translation by making the very least of the refining process. The following discussion shall reveal the steps taken to refine the translation, this means that the process shall be described systematically.

As suggested by the results, the discussion shall describe the process or steps in which the source language is corrected and refined in order to achieve better machine translation. This is important because of there lies and assumptions that the writer had assumed the reader to have understood the information shared, or at least it had been thought that the reader had known it. To provide the reader with the actual data of

the source language, the writer shall expose it in this section.

ABSTRAK

Sejak 2014 telah sering terjadi kebakaran hutan di provinsi Riau, Indonesia yang sebagian besar terjadi karena pembakaran hutan secara ilegal oleh penduduk setempat. Bencana ini telah menyebabkan gangguan pada aspek ekonomi, kesehatan, transportasi, dan lingkungan dan dilaporkan berdampak pada negara-negara tetangga seperti Malaysia dan Singapura. Penelitian ini bertujuan untuk menganalisis program tanggung jawab sosial perusahaan (CSR) di Perusahaan yang memperhatikan bahaya kebakaran lahan dan hutan oleh PT. Riau Andalan Pulp And Paper (PT. RAPP) sebagai strategi untuk mengurangi lahan yang terbakar dengan keterlibatan positif dari masyarakat. Metode penelitian menggunakan penelitian kuantitatif dan kualitatif dengan menggunakan teknik pengambilan sampel non-probabilitas, analisis isi, literasi dokumen yang berasal dari perusahaan. Serta penyebaran kuesioner untuk setiap entitas yang terlibat aktif dengan kegiatan CSR. Penelitian ini mengidentifikasi efektivitas komunikasi CSR, manajemen pesan dan pemilihan media, manajemen program, dan keterlibatan mitra dan proses evaluasi sebagai faktor penentu untuk menangani masalah sosial untuk menciptakan nilai bagi masyarakat. Riset menunjukkan bahwa sebanyak 57,1 % evaluasi program ini memberikan dampak yang positif

dalam memberikan manfaat meningkatkan kesejahteraan kelompok dan komunitas. analisis keseluruhan dari proses penelitian bahwa program CSR 'memiliki kontribusi positif terhadap nilai sosial. Penelitian menggunakan metode

penelitian survey kuantitative untuk mengetahui evaluasi program komunikasi CSR; mengetahui bagaimana pesan utama dalam program FFV, pemilihan media dan evaluasi dampak bagi masyarakat dan bagi reputasi perusahaan.

Kata kunci: Komunikasi CSR, Indonesia, Desa Bebas Api, Reputasi Perusahaan

C. Correcting the Source Language

In this part, every clause is identified for any possible error in the sentence. The errors may come in many forms depending on the writer's linguistic experience; namely their knowledge of the source language. This part is merely dedicated to discovering the way in which Google Translate can easily understand the order through language formation.

*Sejak 2014, telah sering terjadi **kebakaran hutan** di provinsi Riau, Indonesia yang sebagian **besarnya** **dipicu oleh pembakaran hutan** secara ilegal oleh penduduk setempat.*

The correction made in this sentence is to make it a more effective sentence by placing appropriate comma and reducing similar word terjadi and change it with dipicu oleh. In addition, the word kebakaran hutan and pembakaran hutan must be taken into account because both words may have similar equivalent forest fire in English. However, both have different sense of meaning where the former simply means to the general concept of forest fire whereas the later means a forest that is burnt which implies intentional activities by the doer.

*Bencana ini telah menyebabkan gangguan pada aspek ekonomi, kesehatan, transportasi, dan lingkungan, dan **hal ini telah** dilaporkan berdampak pada negara-negara tetangga seperti Malaysia dan Singapura.*

The correction is made by adding the pronoun subjek hal ini which refers back to bencana and the time aspect of telah so that Google Translate can identify which action happens first.

1) Penelitian ini bertujuan untuk menganalisis program tanggung jawab sosial perusahaan (CSR) 2) di Perusahaan yang memperhatikan bahaya kebakaran lahan dan hutan oleh PT. Riau Andalan Pulp And Paper (PT. RAPP) 3) sebagai strategi untuk mengurangi lahan yang terbakar dengan keterlibatan positif dari masyarakat.

There are at least three major clauses suggested by this sentence as indicated by the bold numbers. The idea in the first clause is clear enough, however when connecting it to the second clause, it seems that there is no exact boundary whether a company is doing its CSR in another company or what. This complication is added to the emergence of PT. Riau Andalan Pulp and Paper (PT. RAPP), implying the existence of other companies. However, looking at the main idea, the three companies actually refer back to the same company. Therefore, correction is required for this sentence by modifying the structure, suggested correction:

Penelitian ini bertujuan untuk menganalisis program tanggung jawab sosial perusahaan (CSR) yang programnya memperhatikan bahaya kebakaran lahan dan hutan. Program PT. Riau Andalan Pulp And Paper (PT. RAPP) ini dimaksudkan sebagai strategi

untuk mengurangi lahan yang terbakar melalui keterlibatan positif dari masyarakat.

Next sentence is:

*Penelitian ini mengidentifikasi efektivitas komunikasi CSR, manajemen pesan dan pemilihan media, manajemen program, dan keterlibatan mitra. **Penelitian ini juga mengidentifikasi** proses evaluasi sebagai faktor penentu untuk menangani masalah sosial **guna** menciptakan nilai bagi masyarakat.*

This is a run-on [11] sentence, meaning that the clauses constructing the sentence are embedded inappropriately. Therefore, the correction made is by splitting the sentence into simpler forms as suggested in the bold expression. In addition to that, the subject is made visible for Google Translate to detect. The changing word *untuk* into *guna* is intended to differ the use of *to* in English that is followed by infinitive, whereas in this sense *untuk* is actually intended for in order to.

*Analisis keseluruhan dari proses penelitian **menunjukkan** bahwa program CSR 'memiliki kontribusi positif terhadap nilai sosial.*

This is a fragment [11]; no verb is present here. Therefore, adding *menunjukkan* as the required verb simply fix the problem.

(1) *Metode penelitian menggunakan penelitian kuantitatif dan kualitatif dengan menggunakan teknik pengambilan sampel non-probabilitas, analisis isi, literasi dokumen yang berasal dari perusahaan.* (2) *Serta penyebaran kuesioner untuk setiap entitas yang terlibat aktif dengan kegiatan CSR.*

The case occurs in the sentence (1) is word redundancy in the word *penelitian* and ambiguity in terms of whether the technique used is applicable for both methods, moreover the sentence is using an active sentence which means that metode penelitian is actively doing "the job". In addition, sentence (2) also offers other fragments and ambiguous concept; whether it is part of the method or technique and no verb indicating the main clause. Therefore, correction is required, suggestion:

Metode penelitian yang digunakan adalah kuantitatif dan kualitatif dengan teknik pengambilan sampel non-probabilitas, analisis isi, dan literasi dokumen yang berasal dari perusahaan. Selain itu, penyebaran kuesioner juga dilakukan untuk setiap entitas yang terlibat aktif dalam kegiatan CSR.

The following sentence also offers almost similar case where the research method reappears but only the quantitative one. The use of active voice *menggunakan* implies the same case as the previous one. The following is the correction made to the sentence.

*Metode penelitian survey kuantitative **digunakan** untuk mengetahui evaluasi program komunikasi CSR, mengetahui bagaimana pesan utama dalam program FFV, **serta mengetahui** pemilihan media dan evaluasi dampak bagi masyarakat dan bagi reputasi perusahaan.*

The final result of the correction is as follow:

*Sejak 2014, telah sering terjadi **kebakaran hutan** di provinsi Riau, Indonesia yang sebagian besarnya **dipicu oleh pembakaran hutan** secara ilegal oleh penduduk setempat. Bencana ini telah menyebabkan gangguan*

pada aspek ekonomi, kesehatan, transportasi, dan lingkungan, dan hal ini telah dilaporkan berdampak pada negara-negara tetangga seperti Malaysia dan Singapura. Penelitian ini bertujuan untuk menganalisis program tanggung jawab sosial perusahaan (CSR) yang programnya memperhatikan bahaya kebakaran lahan dan hutan. Program PT. Riau Andalan Pulp And Paper (PT. RAPP) ini dimaksudkan sebagai strategi untuk mengurangi lahan yang terbakar melalui keterlibatan positif dari masyarakat. Penelitian ini mengidentifikasi efektivitas komunikasi CSR, manajemen pesan dan pemilihan media, manajemen program, dan keterlibatan mitra. **Penelitian ini juga mengidentifikasi** proses evaluasi sebagai faktor penentu untuk menangani masalah sosial guna menciptakan nilai bagi masyarakat. Riset menunjukkan bahwa sebanyak 57,1 % evaluasi program ini memberikan dampak yang positif dalam memberikan manfaat meningkatkan kesejahteraan kelompok dan komunitas. Analisis keseluruhan dari proses penelitian **menunjukkan** bahwa program CSR 'memiliki kontribusi positif terhadap nilai sosial. Metode penelitian yang digunakan adalah kuantitatif dan kualitatif dengan teknik pengambilan sampel non-probabilitas, analisis isi, dan literasi dokumen yang berasal dari perusahaan. Selain itu, penyebaran kuesioner juga dilakukan untuk setiap entitas yang terlibat aktif dalam kegiatan CSR. Metode penelitian survey kuantitative **digunakan** untuk mengetahui evaluasi program komunikasi CSR, mengetahui bagaimana pesan utama dalam program FFV, **serta mengetahui** pemilihan media dan evaluasi dampak bagi masyarakat dan bagi reputasi perusahaan.

To illustrate the effectiveness of the correction, below is illustrated the direct result from Google Translate. The first is

the translation before the correction, and the second is after the correction.

1. Since 2014 there have been frequent forest fires in Riau province, Indonesia, which mostly occur due to illegal burning of forests by local residents. This disaster has caused disruption to economic, health, transportation and environmental aspects and has been reported to have an impact on neighboring countries such as Malaysia and Singapore. This study aims to analyze corporate social responsibility (CSR) programs in companies that pay attention to the dangers of land and forest fires by PT. Riau Andalan Pulp and Paper (PT. RAPP) as a strategy to reduce burned land with positive involvement from the community. The research method uses quantitative and qualitative research using non-probability sampling techniques, content analysis, and document literacy from companies. As well as distributing questionnaires to each entity that is actively involved with CSR activities. This study identifies the effectiveness of CSR communications, message management and media selection, program management, and partner involvement and evaluation processes as determining factors for dealing with social problems to create value for society. Research shows that as many as 57.1% of evaluations of this program have a positive impact on providing benefits to improve the welfare of groups and communities. The overall analysis of the research process that CSR programs' have a positive contribution to social value. The research uses a quantitative survey research method to determine the evaluation of CSR

communication programs; find out how the main message in the FFV program, media selection and impact evaluation for the community and for the company's reputation.

2. Since 2014, there have been frequent forest fires in Riau province, Indonesia, most of which were triggered by illegal forest fires by local residents. This disaster has caused disruption to the economic, health, transportation, and environmental aspects, and this has been reported to have affected neighbouring countries such as Malaysia and Singapore. This study aims to analyse corporate social responsibility (CSR) programs whose programs pay attention to the dangers of land and forest fires. PT. Riau Andalan Pulp and Paper (PT. RAPP) program is intended as a strategy to reduce burning land through positive involvement from the community. This research identifies the effectiveness of CSR communication, message management and media selection, program management, and partner involvement. This study also identified the evaluation process as a determining factor for dealing with social problems in order to create value for the community. Research shows that as much as 57.1% evaluation of this program has a positive impact in providing benefits to improve the welfare of groups and communities. The overall analysis of the research process shows that CSR programs' have a positive contribution to social value. The research method used is quantitative and qualitative with non-probability sampling techniques, content analysis, and document literacy from companies. In addition, the distribution of questionnaires was also carried out for

each entity that was actively involved in CSR activities. Quantitative survey research methods are used to find out the evaluation of CSR communication programs, find out how the main messages in the FFV program are, as well as find out media selection and impact evaluation for the community and for the company's reputation.

D. Do a Sentence by Sentence Translation

The second step for refining the machine translation is to make sure that you do not directly copy the whole paragraph. Rather, you make a sentence by sentence translation or phrase by phrase when necessary. This process is required to make sure that you make the least error of the translation and at the same time, it eases you to locate the errors and make the correction on the errors.

The result of this step has already been illustrated in the previous discussion in (2). As it can be seen, the result shows better quality translation than that in (1).

E. Consult Dictionary for Idiomatic Expression

This step addresses the issue in which idiomatic expression may appear. The expression may be culturally bound or is required to the presence in a particular genre; in this term, the scientific expression is preferred rather than the colloquial expression. Therefore, it is important to seek any phrasal words that may degrade the scientific quality of the abstract. From the refined result in (2), there are at least 3 expressions indicated

to serve this issue, as illustrated in the following findings in the bold type:

(1) This study aims to analyse corporate social responsibility (CSR) programs whose programs **pay attention** to the dangers of land and forest fires.

(2) PT. Riau Andalan Pulp and Paper (PT. RAPP) program is intended as a strategy to reduce **burning land** through positive involvement from the community.

(3) Quantitative survey research methods are used to **find out** the evaluation of CSR communication programs, **find out** how the main messages in the FFV program are, as well as **find out** media selection and impact evaluation for the community and for the company's reputation.

The phrasal word *pay attention* may represent the equivalent of *memberi*

perhatian and may also represent the sense intended in this machine translation. However, this use of the phrasal word may reduce the scientific quality and therefore we must seek one term that best represents this term.

This can be done by blocking the expression to *pay attention* to the mouse, right-click on it, and find the menu *synonym* and find thesaurus (see Fig.1). You will then find several options for replacing this expression. Pick one word that is an only single word and consult the dictionary for its use. Is it the same sense as *pay attention* (**concentrate**) or not? If yes, then this may be the most appropriate choice for replacing the expression selected. The same procedure can also be implemented to the expression *find out* (**discover**) as seen in Fig. 2 Below is the illustration:

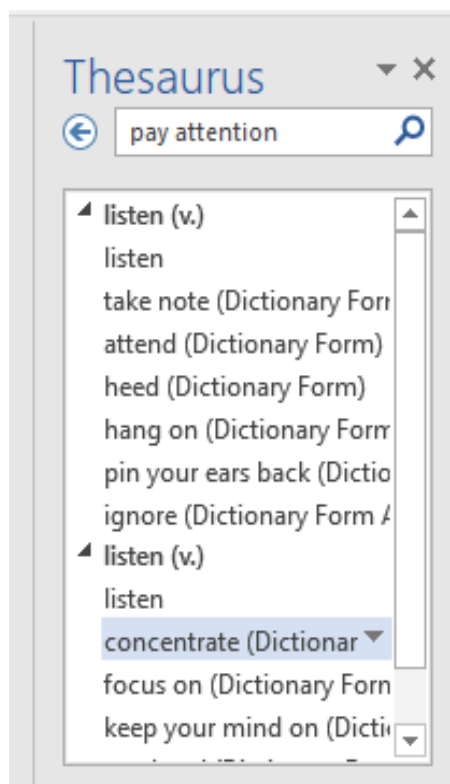


Fig. 1. Finding best equivalent for Pay Attention

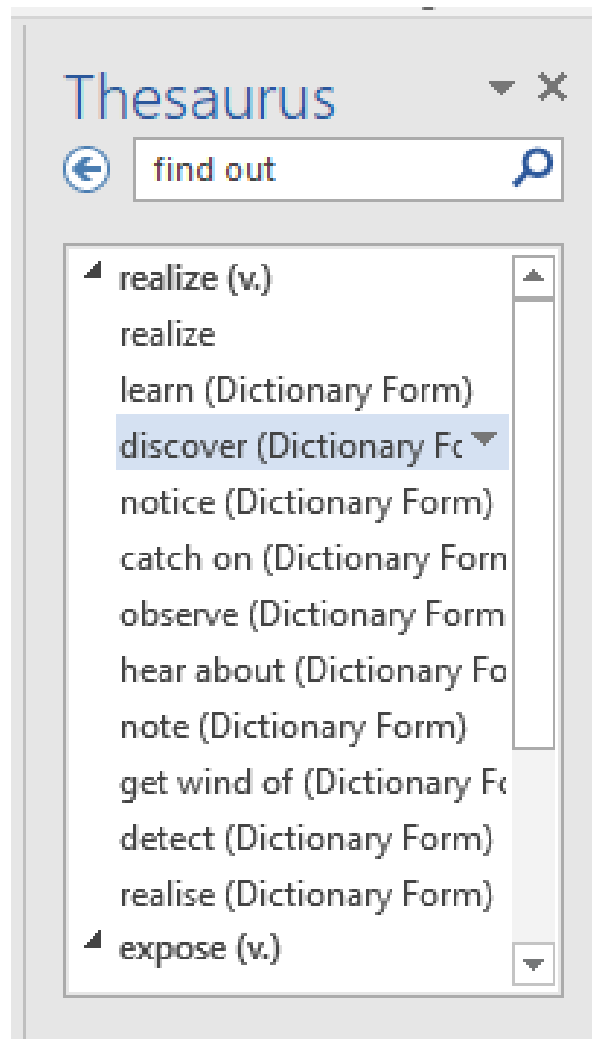


Fig. 2. Finding best equivalent for Find Out

However, you may not find a similar case for burning land (lahan yang terbakar) because it is colloquial. The most natural expression to replace this is by making a proposition. Lahan yang terbakar can be broken down into two propositions: (a) there is a land, and (b) the land is exposed to fire. Through these propositions, we can then make a new expression by involving the expression in the proposition; namely, land, exposed, and fire. From this, we can then change it into fire-exposed land as the substitution to burning land expression.

F. Do a Final Checking to the Translation.

This final step is actually a reverse process of machine translation. After refining the abstract into the following form:

Since 2014, there have been frequent forest fires in Riau province, Indonesia, most of which were triggered by illegal forest fires by local residents. This disaster has caused disruption to the economic, health, transportation, and environmental aspects, and this has been reported to have affected neighbouring countries such as Malaysia and Singapore. This study aims to analyse corporate social responsibility (CSR) programs whose programs **concentrate** on the dangers of land and forest fires.

PT. Riau Andalan Pulp and Paper (PT. RAPP) program is intended as a strategy to reduce **fire-exposed land** through positive involvement from the community. This research identifies the effectiveness of CSR communication, message management and media selection, program management, and partner involvement. This study also identified the evaluation process as a determining factor for dealing with social problems in order to create value for the community. Research shows that as much as 57.1% evaluation of this program has a positive impact on providing benefits to improve the welfare of groups and communities. The overall analysis of the research process shows that CSR programs' have a positive contribution to social value. The research method used is quantitative and qualitative with non-probability sampling techniques, content analysis, and document literacy from companies. In addition, the distribution of questionnaires was also carried out for each entity that was actively involved in CSR activities. Quantitative survey research methods are used to **discover** the evaluation of CSR communication programs, **discover** how the main messages in the FFV program are, as well as discover media selection and impact evaluation for the community and for the company's reputation.

Then this refined machine translation is to be translated back into Bahasa Indonesia. This is to make sure that the translation is as equivalent and natural as the source language. The following is the final result of the final checking:

Sejak 2014, sering terjadi kebakaran hutan di provinsi Riau, Indonesia, yang sebagian besar

dipicu oleh kebakaran hutan ilegal oleh penduduk setempat. Bencana ini telah menyebabkan gangguan pada aspek ekonomi, kesehatan, transportasi, dan lingkungan, dan ini dilaporkan telah mempengaruhi negara-negara tetangga seperti Malaysia dan Singapura. Penelitian ini bertujuan untuk menganalisis program tanggung jawab sosial perusahaan (CSR) yang programnya berkonsentrasi pada bahaya kebakaran lahan dan hutan. PT. Program Pulp dan Kertas Andalan Riau (PT. RAPP) dimaksudkan sebagai strategi untuk mengurangi lahan yang terkena api melalui keterlibatan positif dari masyarakat. Penelitian ini mengidentifikasi efektivitas komunikasi CSR, manajemen pesan dan pemilihan media, manajemen program, dan keterlibatan mitra. Studi ini juga mengidentifikasi proses evaluasi sebagai faktor penentu untuk menangani masalah sosial untuk menciptakan nilai bagi masyarakat. Penelitian menunjukkan bahwa sebanyak 57,1% evaluasi program ini memiliki dampak positif dalam memberikan manfaat untuk meningkatkan kesejahteraan kelompok dan masyarakat. Analisis keseluruhan dari proses penelitian menunjukkan bahwa program CSR 'memiliki kontribusi positif terhadap nilai sosial. Metode penelitian yang digunakan adalah kuantitatif dan kualitatif dengan teknik pengambilan sampel non-probabilitas, analisis konten, dan literasi dokumen dari perusahaan. Selain itu, distribusi kuesioner juga dilakukan untuk setiap entitas yang terlibat aktif dalam kegiatan CSR. Metode penelitian survei kuantitatif digunakan untuk menemukan evaluasi program komunikasi CSR, menemukan bagaimana pesan utama dalam program FFV, serta menemukan pemilihan media dan evaluasi dampak bagi masyarakat dan untuk reputasi perusahaan.

4. CONCLUSION

In this research, refining machine translation involves four steps. Each step requires language analysis in order to achieve the finest form of the target language. This shows that human systematic way of thinking still largely involves in the process. It can also be said that the machine “only” helps them in accelerating the translation process.

ssAcknowledgement

The present we would like to address the highest gratitude to Universitas Komputer Indonesia (UNIKOM) for funding this research to be presented in International Conference on Business, Economic, Social presented in Sciences, and Humanity (ICOBEST) 2019. Likewise, the organizing committee also deserves the same gratitude for their tireless efforts in managing such a conference. Hopefully, this research can extend the way in which computer-aided tool translation is made.

REFERENCES

- Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, 37, 112-121.
- ElShiekh, A. A. A. (2012). Google translate service: transfer of meaning, distortion or simply a new creation? An investigation into the translation process & problems at google. *English Language and Literature Studies*, 2(1), 56.
- Balahur, A., & Turchi, M. (2012, July). Multilingual sentiment analysis using machine translation?. In *Proceedings of the 3rd workshop in computational approaches to subjectivity and sentiment analysis* (pp. 52-60).
- Garcia, I. (2009). Beyond translation memory: Computers and the professional translator. *The Journal of Specialised Translation*, 12(12), 199-214.
- Afshin, H., & Alaeddini, M. A. (2016). A contrastive analysis of machine translation (Google Translate) and human translation: Efficacy in translating verb tense from English to Persian. *Mediterranean Journal of Social Sciences*, 7(4 S2), 40-40.
- Läubli, S., & Orrego-Carmona, D. (2017). When Google Translate is better than some human colleagues, those people are no longer colleagues.
- Ismail, A., & Hartono, R. (2016, May). Errors made in Google Translate in the Indonesian to English translations of news item texts. In *ELT Forum: Journal of English Language Teaching* (Vol. 5, No. 2).
- Aziz, W., Castilho, S., & Specia, L. (2012, May). PET: a Tool for Post-editing and Assessing Machine Translation. In *LREC* (pp. 3982-3987).
- Rahardjo, M. (2017). Studi kasus dalam penelitian kualitatif: konsep dan prosedurnya.

Flick, U. (2018). *An introduction to qualitative research*. sage.

Oshima, A., & Hogue, A. (2007). *Introduction to academic writing* (p. 3).
Pearson/Longman.